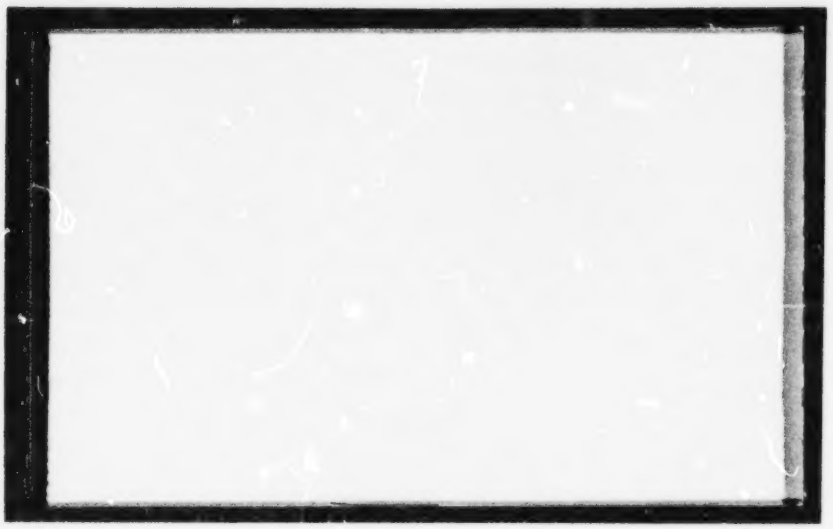
UNIVERSITÉ DE MONTRÉAL



CENTRE DE RECHERCHE SUR LES TRANSPORTS

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

CADRE COMPTABLE RECTANGULAIRE ET ESTIMATION DE MATRICES INPUT-OUTPUT INTERREGIONALES

par

Marc Los

Cette recherche a été soutenue par une subvention du Centre de recherche et de développement de Transports Canada.

Université de Montréal

Centre de recherche sur les transports - Publication # 160
Département de sciences économiques - Cahier No 80-05

Février 1980 - Révisé en août 1980

Dépôt légal - 2e trimestre 1980

RÉSUMÉ

Cet article propose une méthode d'estimation d'une matrice input-output interrégionale à partir d'un cadre comptable rectangulaire tel que celui qui existe pour le Canada. La méthode repose sur l'utilisation des informations statistiques suivantes pour une année de base: matrices régionales des utilisations de biens ou services par les différents secteurs productifs de la région, matrices régionales des productions de biens ou services par les différents secteurs productifs de la région, matrices des exportations internationales de biens et services à partir des régions, matrices d'importations internationales de biens et services par les différentes régions, demandes finales régionalisées pour chaque bien ou service, matrices de flux interrégionaux de biens ou services.

La méthodologie proposée s'appuie d'une part sur une représentation explicite des flux intersectoriels de biens ou services et d'autre part sur la théorie de l'information.

ABSTRACT

This article describes a method for estimating an interregional input-output matrix within a rectangular accounting framework such as it exists for Canada. It is assumed that the following data are known for a base year: regional use matrices, regional output matrices, international exports of goods or services for each region, international imports of goods or services for each region, regional final demands for goods or services, matrices of interregional flows of goods or services. The proposed method relies on an explicit representation of intersectoral flows of goods and services and also on information theory.

I - Introduction

Les matrices input-output interrégionales constituent un outil difficilement remplaçable de modélisation des interdépendances entre plusieurs économies régionales.[1] Au Canada, le Ministère de l'Expansion économique régionale (1976 et 1977) avait construit un tableau input-output interprovincial à un niveau sectoriel agrégé (34 secteurs). Plus récemment, la Division de l'Analyse structurelle de Statistique Canada a mis au point, à partir de données d'enquêtes pour 1974, un modèle input-output interprovincial pour le Canada, à un niveau de désagrégation considérable pour les secteurs (200) comme pour les biens et services (600).[2]

En vue de différentes utilisations, telles que l'utilisation répétée du modèle dans des simulations, la mise à jour pour une année ultérieure à 1974, le calcul de certains coefficients entrant dans des modèles économiques autres que le modèle interprovincial, etc... on peut être amené à vouloir utiliser les données input-output régionalisées de Statistique Canada à un niveau plus agrégé (spatial, sectoriel, pour les biens ou services) pour des raisons pratiques de calcul comme pour des raisons de confidentialité, afin de construire un modèle input-output interrégional "pur" (Isard, 1951, 1953). Ce dernier se

---

[1] Cf. Isard (1951, 1953), Leontieff (1953), Chenery, Clark et Cao Pinna (1953), Moses (1955) pour les articles pionniers dans le domaine de l'input-output interrégional ainsi que Riefler (1973) pour une revue de l'état de l'art et une bonne bibliographie pour la période allant jusqu'à 1973.

[2] Cf Hoffman et Kent (1976), Hoffman (1978), Gaston (1979) et le chapitre 7 de Statistique Canada, Division de l'Analyse structurelle (1976) pour la conception et la compilation du modèle interprovincial de Statistique Canada.

traduit par une matrice de coefficients structurels intersectoriels, interrégionaux, conçue comme une simple désagrégation spatiale du modèle input-output de Leontieff.[1]

Cet article propose une méthode d'estimation d'une telle matrice input-output interrégionale permettant de prendre en compte les informations statistiques régionalisées suivantes, disponibles auprès de Statistique Canada à un niveau moins détaillé que celui auquel il a été compilé: matrices régionales d'inputs, matrices régionales d'outputs, matrices de flux interrégionaux de biens et services, matrices des exportations internationales de biens et services à partir des régions, matrices d'importations internationales de biens et services pour les différentes régions, demandes finales régionales pour chaque bien ou service. Comme il n'existe pas de données disponibles *directement* sur les flux intersectoriels et interrégionaux, mais seulement les informations indiquées précédemment, le problème d'estimation consiste à *inférer* une matrice de coefficients compatible avec des contraintes traduisant les données disponibles.

Plusieurs approches sont à priori possibles. Une approche souvent utilisée est celle dite de "Chenery-Moses" (Chenery *et al.*, 1953; Moses, 1955) ou à coefficients-colonnes constants. Cette approche classique fait l'hypothèse que tous les secteurs d'une région donnée

---

[1] Cf. Courville *et al.* (1979) et Van Peeterssen (1979) pour la construction et l'utilisation d'une telle matrice afin d'étudier les relations du Québec avec l'Ontario et le reste du Canada.

ont la même répartition spatiale de leurs inputs. Cela permet de construire la matrice cherchée uniquement à partir d'informations sur le commerce interprovincial par bien ou service. Cette approche a été utilisée par Van Peeterssen (1979), ainsi d'ailleurs que dans le modèle interprovincial de Statistique Canada (Statistique Canada, 1976, chapitre 7).

L'hypothèse de Chenery-Moses est plutôt restrictive et n'a pas de justification théorique. De plus, elle oblige à laisser tomber dans l'estimation une partie des informations statistiques disponibles (en particulier les exportations internationales en provenance des régions et les demandes finales régionales en biens ou services) et ne s'adapte pas facilement à la prise en compte d'informations supplémentaires, même partielles.

Cet article propose une alternative à la méthode de Chenery-Moses. La méthodologie alternative proposée consiste à considérer le problème d'estimation d'une matrice input-output interrégionale directement comme un problème d'*inférence statistique*: il s'agit de trouver la matrice qui satisfait à l'information dont on dispose tout en optimisant un critère statistique. La méthodologie proposée s'inspire directement du cadre conceptuel proposé par S.M. Macgill (1977,1978) lequel repose sur une représentation explicite des flux intersectoriels de biens et services et s'adapte bien au cadre comptable rectangulaire utilisé au Canada (Statistique Canada, 1976; Gigantes, 1970). Dans ce cadre conceptuel, le problème d'estimation posé se ramène à inférer une distribution de probabilité qui obéit à un certain nombre de contraintes.

De tels problèmes d'inférence peuvent se résoudre en utilisant le formalisme de la maximisation de l'entropie qui a un fondement rigoureux en théorie de l'information[1]: la fonction entropie est la seule mesure d'incertitude d'une distribution de probabilité qui obéit à un certain nombre d'axiomes raisonnables.

Dans cet article nous utiliserons le principe de maximisation de l'entropie dans l'estimation de la matrice input-output interrégionale. L'algorithme proposé pour résoudre le programme mathématique résultant de la méthodologie suggérée peut être considéré comme une extension et adaptation de la méthode RAS, bien qu'il permette d'échapper à certaines critiques qui ont été faites de cette méthode.

Le reste de l'article est organisé comme suit: la section II formule la méthode d'estimation suggérée comme problème de programmation mathématique. La section III décrit un algorithme pour résoudre ce programme ainsi que la manière de calculer les coefficients structurels interrégionaux intersectoriels et les coefficients d'importation pour chaque secteur dans chaque région. La section IV propose certaines extensions et conclut l'article.

---

[1] Cf. Jaynes (1957,1979), Snickars et Weibull (1977), Snickars (1979), Theil (1967), Wilson (1970).

II - Méthode d'estimation

Cette section est consacrée à la présentation d'une méthode d'estimation de coefficients input-output interrégionaux intersectoriels et de coefficients d'importation à partir d'informations input-output régionales comme celles qui sont disponibles auprès de Statistique Canada.

Les notations suivantes seront utilisées:

i,j=1,...R	représentent les régions
i,j=0	représentent le reste du monde
m,n=1,...,N-1	représentent les secteurs productifs de l'économie
n=N	représente le secteur de la demande finale
k=1,...,K	représentent les biens et services produits par les secteurs productifs.

Un même bien k peut être produit par plusieurs secteurs de l'économie et réciproquement, un même secteur m peut produire différents biens ou services k. Il s'agit d'estimer un ensemble de coefficients $\{b_{ij}^{mn}/i,j=1,\ldots,R;\ m,n=1,\ldots,N-1\}$ où b_{ij}^{mn} représente la quantité d'output du secteur m dans la région i utilisée comme input par unité d'output du secteur n dans la région j. Par ailleurs on pourra aussi estimer des coefficients $\{\mu_j^n/j=1,\ldots,R;\ n=1,\ldots,N-1\}$ représentant la valeur des importations internationales du secteur n dans la région j par unité d'output de ce secteur. Le résultat de l'estimation sera donc un modèle input-output interrégional "pur" (Riefler, 1973) où on ne s'intéresse qu'aux flux *intersectoriels* et non aux flux de biens et services.

On suppose connues pour une année de base les informations statistiques suivantes (Gaston, 1979) :

U_j^{nk}, j=1,...,R; n=1,...,N; k=1,...,K

représente l'utilisation de bien k par le secteur n (demande finale incluse) dans la région j.

V_i^{mk}, i=1,...,R; m=1,...,N-1; k=1,...,K

représente la production de bien k par le secteur m dans la région i.

X_i^k, i=1,...,R; k=1,...,K

représente les exportations internationales de bien k par l'ensemble des secteurs productifs de la région i.

M_j^k, j=1,...,R; k=1,...,K

représente les importations internationales de bien k par l'ensemble des secteurs (demande finale incluse) de la région consommatrice j.

Z_{ij}^k, i=1,...,R; j=1,...,R; k=1,...,K

représente le commerce total de bien k de la région i vers la région j, quel que soit le secteur d'origine ou le secteur de destination

Y_j^{ny}, j=1,...,R; n=1,...,N-1; y=1,...,W

représente l'utilisation du facteur primaire y par le secteur productif n dans la région j.

L'idée essentielle de l'approche consiste à estimer pour l'année de base des flux $\{u_{ij}^{mnk}/i=1,...,R; j=1,...,R; m=1,...,N-1; n=1,...,N; k=1,...,K\}$ où u_{ij}^{mnk} représente le flux de bien k en provenance du secteur m de la région i et destiné au secteur n de la région j; ainsi que des flux $\{u_{io}^{mnk}/i=1,...,R; m=1,...,N-1; n=1,...,N; k=1,...,K\}$ pour les exportations internationales et des flux $\{u_{oj}^{mnk}/j=1,...,R; m=1,...,N-1; n=1,...,N; k=1,...,K\}$

pour les importations internationales. Il faut noter dès maintenant le fait qu'à aucun moment il ne sera nécessaire de stocker de telles grandeurs en mémoire, ce qui serait impossible en pratique.[1] Ce point sera développé en détail dans la section suivante.

Pour être cohérentes les données doivent vérifier pour chaque bien et chaque région, les relations suivantes:

$$\sum_{m=1}^{N-1} V_i^{mk} + M_i^k + \sum_{j=1}^{R} Z_{ji}^k = \sum_{n=1}^{N} U_i^{nk} + x_i^k + \sum_{j=1}^{R} Z_{ij}^k \qquad \begin{matrix} i=1,\ldots,R \\ k=1,\ldots,K \end{matrix} \qquad (1)$$

Production + importations = utilisation + exportations. Elles doivent aussi vérifier pour chaque secteur productif les relations:

$$\sum_{k=1}^{K} V_j^{nk} - \sum_{y=1}^{W} Y_j^{ny} = \sum_{k=1}^{K} U_j^{nk} \qquad \begin{matrix} j=1,\ldots,R \\ n=1,\ldots,N-1 \end{matrix} \qquad (2)$$

Valeur de l'output-dépenses en facteurs primaires = dépenses en biens intermédiaires.

La méthode d'estimation proposée consiste à déterminer les flux u_{ij}^{mnk}, u_{io}^{mnk}, u_{oj}^{mnk} qui forment la solution du programme mathématique suivant:

$$\text{Min} \sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N}\sum_{k=1}^{K} u_{ij}^{mnk} \ln u_{ij}^{mnk} + \sum_{i=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N}\sum_{k=1}^{K} u_{io}^{mnk} \ln u_{io}^{mnk} + \sum_{j=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N}\sum_{k=1}^{K} u_{oj}^{mnk} \ln u_{oj}^{mnk} \qquad (3)$$

[1] Par exemple, avec le nombre de régions (3), de secteurs (44) et de biens (83) utilisés par Courville *et al.* (1979) cela ferait environ 1 400 000 nombres à stocker.

t.q.

$$\sum_{n=1}^{N} \sum_{j=0}^{R} u_{ij}^{mnk} = V_i^{mk} \qquad \begin{matrix} i=1,\ldots,R \\ m=1,\ldots,N-1 \\ k=1,\ldots,K \end{matrix} \qquad (4)$$

$$\sum_{m=1}^{N-1} \sum_{i=0}^{R} u_{ij}^{mnk} = U_j^{nk} \qquad \begin{matrix} j=1,\ldots,R \\ n=1,\ldots,N \\ k=1,\ldots,K \end{matrix} \qquad (5)$$

$$\sum_{m=1}^{N-1} \sum_{n=1}^{N} u_{io}^{mnk} = X_i^k \qquad \begin{matrix} i=1,\ldots,R \\ k=1,\ldots,K \end{matrix} \qquad (6)$$

$$\sum_{m=1}^{N-1} \sum_{n=1}^{N} u_{oj}^{mnk} = M_j^k \qquad \begin{matrix} j=1,\ldots,R \\ k=1,\ldots,K \end{matrix} \qquad (7)$$

$$\sum_{m=1}^{N-1} \sum_{n=1}^{N} u_{ij}^{mnk} = Z_{ij}^k \qquad \begin{matrix} i,j=1,\ldots,R \\ k=1,\ldots,K \end{matrix} \qquad (8)$$

$$u_{ij}^{mnk} \geq 0 \qquad \begin{matrix} i,j=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \\ k=1,\ldots,K \end{matrix} \qquad (9)$$

$$u_{io}^{mnk} \geq 0 \qquad \begin{matrix} i=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \\ k=1,\ldots,K \end{matrix} \qquad (10)$$

$$u_{oj}^{mnk} \geq 0 \qquad \begin{matrix} j=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \\ k=1,\ldots,K \end{matrix} \qquad (11)$$

Les contraintes (4), (5), (6), (7) et (8) expriment simplement que certains totaux obtenus à partir des flux élémentaires u_{ij}^{mnk}, u_{io}^{mnk} ou u_{oj}^{mnk} doivent égaler les quantités connues: V_i^{mk}, U_j^{nk}, X_i^k, M_j^k, Z_{ij}^k. Elles explicitent donc toute l'information statistique disponible. Parmi tous les flux satisfaisant à ces contraintes, le problème d'estimation consiste à choisir

ceux qui maximisent l'entropie d'une certaine distribution de probabilité[1], celle dont les probabilités sont

$$\frac{u_{ij}^{mnk}}{T}, \quad \frac{u_{io}^{mnk}}{T}, \quad \frac{u_{oj}^{mnk}}{T}$$

$$\text{avec } T = \sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N}\sum_{k=1}^{K} u_{ij}^{mnk} + \sum_{i=1}^{R}\sum_{m=1}^{N}\sum_{n=1}^{N}\sum_{k=1}^{K} u_{io}^{mnk} + \sum_{j=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N}\sum_{k=1}^{K} u_{oj}^{mnk} \tag{12}$$

Il est facile de vérifier que minimiser la fonction objectif (3) est équivalent à maximiser l'entropie de la distribution de probabilité précédente si T est connue. Il en est bien ainsi puisque la somme des transactions de l'économie est implicite dans les données, et est connue séparément pour chaque bien:

$$T^k = \sum_{i=1}^{R}\sum_{j=1}^{R} Z_{ij}^k + \sum_{i=1}^{R} X_i^k + \sum_{j=1}^{R} M_j^k \qquad k=1,\ldots,K \tag{13}$$

$$T = \sum_{k=1}^{K} T^k \tag{14}$$

Les coefficients b_{ij}^{mn} et μ_j^n sont calculables à partir de la solution du programme mathématique précédent:

$$b_{ij}^{mn} = \frac{\sum_{k=1}^{K} u_{ij}^{mnk}}{\sum_{k=1}^{K} V_j^{nk}} \qquad \begin{array}{l} i\ 1,\ldots,R \\ j\ 1,\ldots,R \\ m\ 1,\ldots,N\text{-}1 \\ n\ 1,\ldots,N\text{-}1 \end{array} \tag{15}$$

$$\mu_j^n = \frac{\sum_{m=1}^{N-1}\sum_{k=1}^{K} u_{oj}^{mnk}}{\sum_{k=1}^{K} V_j^{nk}} \qquad \begin{array}{l} j=1,\ldots,R \\ n=1,\ldots,N\text{-}1 \end{array} \tag{16}$$

[1] Ceci revient intuitivement à déterminer les flux les plus probables, compatibles avec les données (voir Jaynes, 1957; W lson, 1970; Snickars et Weibull, 1977).

La section suivante indiquera comment résoudre ce programme mathématique en pratique.

III - Algorithme

Il faut noter tout d'abord que la fonction objectif et toutes les contraintes sont séparables en autant de termes qu'il y a de biens ou services k et que le total des transactions en bien k, T^k est connu et résulte des contraintes (6), (7) et (8). Par conséquent le programme (3) - (11) est décomposable en autant de sous-problèmes qu'il y a de biens k. De plus, il apparaîtra que les contraintes de non-négativité sont redondantes. Nous sommes donc amenés à résoudre K problèmes du type suivant où l'indice k a été éliminé pour simplifier les notations.

$$\text{Min} \sum_{i=1}^{R}\sum_{j=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N} u_{ij}^{mn} \ln u_{ij}^{mn} + \sum_{i=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N} u_{io}^{mn} \ln u_{io}^{mn} + \sum_{j=1}^{R}\sum_{m=1}^{N-1}\sum_{n=1}^{N} u_{oj}^{mn} \ln u_{oj}^{mn} \tag{17}$$

t.q.

$$\sum_{n=1}^{N}\sum_{j=0}^{R} u_{ij}^{mn} = V_i^m \qquad \begin{matrix} i=1,\ldots,R \\ m=1,\ldots,N-1 \end{matrix} \tag{18}$$

$$\sum_{m=1}^{N-1}\sum_{i=0}^{R} u_{ij}^{mn} = U_j^n \qquad \begin{matrix} j=1,\ldots,R \\ n=1,\ldots,N \end{matrix} \tag{19}$$

$$\sum_{m=1}^{N-1}\sum_{n=1}^{N} u_{io}^{mn} = X_i \qquad i=1,\ldots,R \tag{20}$$

$$\sum_{m=1}^{N-1}\sum_{n=1}^{N} u_{oj}^{mn} = M_j \qquad j=1,\ldots,R \tag{21}$$

$$\sum_{m=1}^{N-1}\sum_{n=1}^{N} u_{ij}^{mn} = Z_{ij} \qquad \begin{matrix} i=1,\ldots,R \\ j=1,\ldots,R \end{matrix} \tag{22}$$

Le problème consiste à minimiser une fonction strictement convexe sous contraintes linéaires. Par conséquent les conditions d'optimalité du

premier ordre sont nécessaires et suffisantes pour un optimum et celui-ci est unique. Soient β_i^m, γ_j^n, δ_i, ε_j et ϕ_{ij} les variables duales associées respectivement aux contraintes (18), (19), (20), (21) et (22). On a alors[1]:

$$u_{ij}^{mn} = \exp\{-1-\beta_i^m-\gamma_j^n-\phi_{ij}\} \quad , \qquad \begin{matrix} i,j=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \end{matrix} \qquad (23)$$

$$u_{io}^{mn} = \exp\{-1-\beta_i^m-\delta_i\} \qquad \begin{matrix} i=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \end{matrix} \qquad (24)$$

$$u_{oj}^{mn} = \exp\{-1-\gamma_j^n-\varepsilon_j\} \qquad \begin{matrix} j=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \end{matrix} \qquad (25)$$

Posons

$$B_i^m = \exp\,(-\beta_i^m) \qquad \begin{matrix} i=1,\ldots,R \\ m=1,\ldots,N-1 \end{matrix} \qquad (26)$$

$$C_j^n = \exp\,(-\gamma_j^n) \qquad \begin{matrix} j=1,\ldots,R \\ n=1,\ldots,N \end{matrix} \qquad (27)$$

$$D_i = \exp(-1-\delta_i) \qquad i=1,\ldots,R \qquad (28)$$

$$E_j = \exp(-1-\varepsilon_j) \qquad j=1,\ldots,R \qquad (29)$$

$$F_{ij} = \exp(-1-\phi_{ij}) \qquad i,j=1,\ldots,R \qquad (30)$$

On peut réécrire (23), (24), (25) sous la forme suivante:

$$u_{ij}^{mn} = B_i^m\, C_j^n\, F_{ij} \qquad \begin{matrix} i,j=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \end{matrix} \qquad (31)$$

$$u_{io}^{mn} = B_i^m\, D_i \qquad \begin{matrix} i=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \end{matrix} \qquad (32)$$

$$u_{oj}^{mn} = C_j^n\, E_j \qquad \begin{matrix} j=1,\ldots,R \\ n=1,\ldots,N \end{matrix} \qquad (33)$$

---

[1] On vérifie en même temps que les flux sont tous strictement positifs et donc que les contraintes de non-négativité sont redondantes.

L'algorithme proposé ci-dessous calcule les "facteurs de balancement" B_i^m, C_j^n, F_{ij}, D_i et E_j de manière itérative, afin de satisfaire simultanément aux conditions du premier ordre, (31) - (33), sur les variables duales et aux contraintes, (18) - (22), sur les variables primales.

Supposons données des valeurs initiales quelconques aux facteurs de balancement, soient, $\overline{C}_j^n$, $\overline{F}_{ij}$, $\overline{D}_i$, $\overline{E}_j$.

De (18) on tire:

$$V_i^m = \sum_{n=1}^{N} \sum_{j=1}^{R} B_i^m \overline{C}_j^n \overline{F}_{ij} + \sum_{n=1}^{N} B_i^m \overline{D}_i \tag{34}$$

$$\overline{B}_i^m = \frac{V_i^m}{\sum_{n=1}^{N} \sum_{j=1}^{R} \overline{C}_j^n \overline{F}_{ij} + \sum_{n=1}^{N} \overline{D}_i} \qquad \begin{matrix} i=1,\ldots,R \\ m=1,\ldots,N-1 \end{matrix} \tag{35}$$

On peut alors recalculer les facteurs D_i à partir de (20):

$$X_i = \sum_{m=1}^{N-1} \sum_{n=1}^{N} \overline{B}_i^m D_i \tag{36}$$

$$\overline{D}_i = \frac{X_i}{N \sum_{m=1}^{N-1} \overline{B}_i^m} \qquad i=1,\ldots,R \tag{37}$$

On itère sur (35) et (37) jusqu'à ce que les facteurs $\overline{B}_i^m$ et $\overline{D}_i$ obtenus soient compatibles.

De (19) on tire:

$$U_j^n = \sum_{m=1}^{N-1} \sum_{i=1}^{R} \overline{B}_i^m C_j^n \overline{F}_{ij} + \sum_{m=1}^{N-1} C_j^n \overline{E}_j \tag{38}$$

$$\overline{C}_j^n = \frac{U_j^n}{\sum_{m=1}^{N-1} \sum_{i=1}^{R} \overline{B}_i^m \overline{F}_{ij} + \sum_{m=1}^{N-1} \overline{E}_j} \qquad \begin{matrix} j=1\ldots,R \\ n=1\ldots,N \end{matrix} \tag{39}$$

On peut alors recalculer les facteurs E_j à partir de (21):

$$M_j = \sum_{m=1}^{N-1} \sum_{n=1}^{N} \overline{C}_j^n E_j \tag{40}$$

$$\overline{E}_j = \frac{M_j}{(N-1) \sum_{n=1}^{N} \overline{C}_j^n} \qquad j=1,\ldots,R \tag{41}$$

On itère sur (39) et (41) jusqu'à ce que les facteurs $\overline{C}_j^n$ et $\overline{E}_j$ obtenus soient compatibles.

On calcule alors les facteurs F_{ij} à partir des contraintes (22):

$$Z_{ij} = \sum_{m=1}^{N-1} \sum_{n=1}^{N} \overline{B}_i^m \overline{C}_j^n F_{ij} \tag{42}$$

$$\overline{F}_{ij} = \frac{Z_{ij}}{\sum_{m=1}^{N-1} \sum_{n=1}^{N} \overline{B}_i^m \overline{C}_j^n} \qquad i,j=1,\ldots,R \tag{43}$$

En résumé l'algorithme fonctionne comme suit. On calcule $\overline{B}_i^m$ et $\overline{D}_i$ par balancement sur (35) et (37) les autres facteurs étant fixés, puis $\overline{C}_j^n$ et $\overline{E}_j$ par balancement sur (39) et (41), les autres facteurs étant fixés, puis les facteurs $\overline{F}_{ij}$ à partir des contraintes (43). On continue jusqu'à ce qu'il y ait convergence. Il est clair que, s'il y a convergence, on a obtenu la solution unique du problème d'optimisation, puisqu'on a satisfait aux conditions nécessaires et suffisantes d'optimalité du premier ordre. De plus, il est possible de montrer que les différentes étapes de l'algorithme de balancement proposé consistent à augmenter la fonction objectif du dual du problème (17) - (22) en modifiant successivement chaque variable duale, les autres étant fixées. Comme la fonction objectif est

bornée supérieurement par l'optimum du primal, l'algorithme converge nécessairement et d'après ce qui précède on a alors obtenu l'optimum. On voit ainsi que l'algorithme de balancement n'est qu'une méthode particulière de résolution du problème dual et que d'autres algorithmes de maximisation d'une fonction concave non contrainte pourraient être utilisés.

Il résulte de ce qui précède que les flux u_{ij}^{mnk} obtenus par cette méthode sont donnés par

$$u_{ij}^{mnk} = \chi_{ij}^{k} \, V_i^{mk} \, U_j^{nk} \, Z_{ij}^{k} \qquad \begin{matrix} i,j=1,\ldots,R \\ m=1,\ldots,N-1 \\ n=1,\ldots,N \\ k=1,\ldots,K \end{matrix} \tag{44}$$

où χ_{ij}^{k} est un coefficient de proportionnalité qui dépend de l'ensemble des facteurs de balancement.[1] On a de même

$$u_{io}^{mnk} = \chi_i^{'k} V_i^{mk} \, \chi_i^{k} \tag{45}$$

$$u_{oj}^{mnk} = \chi_j^{''k} \, U_j^{nk} M_j^{k} \tag{46}$$

où $\chi_i^{'k}$ et $\chi_j^{''k}$ sont des facteurs de proportionnalité dépendant de certains facteurs de balancement.

On a

$$\chi_{ij}^{k} = \frac{1}{\sum_{n}^{N} \sum_{j=1}^{R} \bar{C}_j^{nk} \bar{F}_{ij}^{k} + \sum_{n=1}^{N} \bar{D}_i^{k}} \times \frac{1}{\sum_{m=1}^{N-1} \sum_{i=1}^{R} \bar{B}_i^{mk} \bar{F}_{ij}^{k} + \sum_{m=1}^{N-1} \bar{E}_j^{k}} \times \frac{1}{\sum_{m=1}^{N-1} \sum_{n=1}^{N} \bar{B}_i^{mk} \bar{C}_j^{nk}} \tag{47}$$

$$\chi_{ik}^{'} = \frac{1}{\sum_{n=1}^{N} \sum_{j=1}^{R} \bar{C}_j^{nk} \bar{F}_{ij}^{k} + \sum_{n=1}^{N} \bar{D}_i^{k}} \times \frac{1}{N \sum_{m=1}^{N-1} \bar{B}_i^{mk}} \tag{48}$$

---

[1] L'expression (44) est à rapprocher de l'expression analogue (3) de Van Peeterssen (1979) où le coefficient de proportionnalité a une forme beaucoup plus simple.

$$x_j^{*k} = \frac{1}{\sum_{m=1}^{N-1} \sum_{i=1}^{R} \bar{B}_i^{mk} \bar{F}_{ij}^{k} + \sum_{m=1}^{N-1} \bar{E}_j^{k}} \times \frac{1}{(N-1) \sum_{n=1}^{N} \bar{C}_j^{nk}} \qquad (49)$$

Une fois les facteurs de balancement B_i^{mk}, C_j^{nk}, F_{ij}^k, D_i^k, E_j^k, calculés pour chaque bien k, ces grandeurs sont stockées en mémoire et utilisées directement pour le calcul des coefficients input-output interrégionaux et des coefficients d'importation.[1] On a:

$$b_{ij}^{mn} = \frac{\sum_{k=1}^{K} B_i^{mk} C_j^{nk} F_{ij}^k}{\sum_{k=1}^{K} V_j^{nk}} \qquad \begin{matrix} i,j=1,\ldots,R \\ m,n=1,\ldots,N-1 \end{matrix} \qquad (50)$$

$$\mu_j^n = \frac{\sum_{m=1}^{N-1} \sum_{k=1}^{K} C_j^{nk} E_j^k}{\sum_{k=1}^{K} V_j^{nk}} \qquad \begin{matrix} j=1,\ldots,R \\ n=1,\ldots,N-1 \end{matrix} \qquad (51)$$

Par conséquent, il n'est pas nécessaire de stocker en mémoire les flux $\{u_{ij}^{mnk}\}$ pour l'année de base, bien que la méthode d'estimation repose sur une considération explicite de ces flux. On n'a jamais besoin de stocker en mémoire des tableaux de dimensions supérieures à ceux des données ou à ceux que l'on cherche à estimer.

IV - Discussion et conclusions

Cet article a présenté dans le détail une méthode d'estimation d'une matrice input-output interrégionale à partir de matrices régionales d'inputs, de matrices régionales d'outputs, de matrices de flux interrégionaux de biens et services, des exportations et importations internationales de biens et services pour chaque région et de la demande finale par

[1] On pourrait aussi calculer des coefficients d'importation *sectoriels* (Los, 1980).

bien dans chaque région. Contrairement à des méthodes alternatives, telles la méthode de Chenery-Moses, celle qui est proposée dans cet article ne fait aucune hypothèse économique a priori mais cherche à inférer les flux intersectoriels de biens des informations statistiques connues directement.

La méthode d'inférence utilisée (maximisation de l'entropie) a un fondement logique rigoureux et se situe dans la tradition qui interprète les probabilités d'une manière subjective (voir Jaynes, 1979). Elle permet de prendre en compte *toute* l'information disponible mais *rien que* l'information disponible sans apporter de biais artificiel dans l'estimation.

La méthodologie suggérée est suffisamment flexible pour prendre en compte toute information supplémentaire qui serait disponible en plus de celle qui est prise en compte dans les contraintes (4) - (8). Par exemple si l'on connaît en détail la structure géographique des importations d'un secteur donné dans une région donnée on doit en tenir compte en ajoutant des contraintes appropriées. Même des informations plus faibles peuvent être prises en compte; par exemple on peut connaître pour certains secteurs régionaux la part des inputs en provenance de la région même, sans nécessairement connaître la répartition géographique des inputs en provenance de l'extérieur de la région.

Au contraire, si certaines données ne sont connues qu'à un niveau plus agrégé que ce qui est supposé dans les contraintes (4) - (8) on peut adapter la méthode à ce cas. Par exemple si les demandes finales régionales ne sont connues qu'au niveau de l'ensemble des biens, on remplace

l'ensemble des contraintes (5) pertinentes par des contraintes agrégées du type

$$\sum_{k=1}^{K} \sum_{m=1}^{N-1} \sum_{i=0}^{R} u_{ij}^{mNk} = U_j^{N*} \tag{52}$$

où U_j^{N*} représente la demande finale totale de la région j. De même, si pour certains secteurs les utilisations ou les productions ne sont connues qu'au niveau national, on remplace les contraintes (4) ou (5) associées à ces secteurs m_1 et n_1 par des contraintes "nationales" du type:

$$\sum_{i=1}^{R} \sum_{n=1}^{N} \sum_{j=0}^{R} u_{ij}^{m_1 nk} = V_*^{m_1 k} \tag{53}$$

$$\sum_{j=1}^{R} \sum_{m=1}^{N-1} \sum_{i=0}^{R} u_{ij}^{mn_1 k} = U_*^{n_1 k} \tag{54}$$

où $V_*^{m_1 k}$ représente la production de bien k par le secteur m_1 au niveau national et $U_*^{n_1 k}$ représente l'utilisation de bien k par le secteur n_1 au niveau national.

L'algorithme de balancement décrit plus haut doit être adapté à chaque ensemble de contraintes. Si certaines informations ne peuvent être prises en compte que sous forme d'*inégalités* ou sous formes d'*intervalles*, on peut adapter les algorithmes de balancement à ces cas (Ducharme, 1980). L'expérience acquise avec les algorithmes de balancement dans le domaine du transport urbain (voir Ducharme, 1980, par exemple) suggère qu'ils sont extrêmement efficaces.

Il est bon de noter aussi que lorsque l'on dispose d'informations additionnelles à celles qu'utilise l'approche de Chenery-Moses et que

l'on a des raisons économiques de faire confiance a priori à l'hypothèse de Chenery-Moses on peut remplacer la maximisation de l'entropie par la minimisation d'une fonction d'*information*[1] avec des probabilités a priori qui seraient obtenues à partir de l'hypothèse de Chenery-Moses. Dans ce cas la méthode d'inférence consiste à trouver les flux les plus proches des flux "Chenery-Moses" qui satisfont à toutes les informations *supplémentaires* disponibles.

En résumé, on peut se trouver dans les trois cas suivants en ce qui concerne l'information dont on dispose: moins, autant ou plus d'information que celle qui correspond à l'utilisation de la méthode de Chenery-Moses. Si l'on est dans le premier cas, on peut utiliser directement la méthode de maximisation de l'entropie, mais la région réalisable est probablement trop grande pour qu'on puisse attacher beaucoup de confiance à l'estimation obtenue. Si l'on est dans le deuxième cas, on a le choix entre la maximisation de l'entropie et la méthode de Chenery-Moses. Dans le troisième cas on a le choix entre la maximisation de l'entropie et la minimisation de l'information avec des probabilités a priori provenant de la méthode de Chenery-Moses. Il est clair que dans ce troisième cas les estimations obtenues par les deux méthodes alternatives sont d'autant plus proches que la région réalisable qui traduit les informations disponibles est plus petite.

---

[1] La maximisation de l'entropie est un cas particulier de la minimisation de l'information, que l'on obtient quand les probabilités a priori forment une distribution uniforme (voir Snickars et Weibull, 1977).

REMERCIEMENTS

Je remercie Marcel Dagenais, Alain Van Peeterssen et André Martens pour leurs commentaires sur une version antérieure de cet article.

BIBLIOGRAPHIE

CHENERY, H., CLARK, P.G. et CAO PINNA, V. (1953). The structure and growth of the Italian Economy. Rome: U.S. Mutual Security Agency.

COURVILLE, L., DAGENAIS, M., NAPPI, C., VAN PEETERSSEN, A. (1979). La sensibilité des industries au commerce interrégional: le cas du Québec, de l'Ontario et du reste du Canada. Gouvernement du Québec, Ministère des Affaires intergouvernementales.

DUCHARME, R. (1980), Modèles de distribution avec contraintes d'inégalité: calibration et résolution. Centre de recherche sur les transports, Université de Montréal, Publication # 170.

GASTON, C. (1979). Canadian Interregional Input-Output Tables - Sources and Methodology. Statistique Canada, Division de l'Analyse Structurelle.

GIGANTES, T. (1970). The representation of technology in input-output systems, dans Carter, A.P. et Brody, A., eds., Contribution to Input-Output Analysis, North-Holland.

HOFFMAN, R.B. and KENT, J.N. (1976). Design for a commodity by industry interregional input-output model" dans Advances in Input-Output Analysis, K.R. Polenske et J.B. Skolka, eds., Ballinger Publishing Co., Cambridge, Mass, U.S.A.

HOFFMAN, R.B. (1978). Methodology for Compiling Interprovincial Trade Statistics. Working Paper 78-07-27, Division de l'Analyse structurelle, Statistique Canada.

ISARD, W. (1951). Interregional and regional input-output analysis: A model of a space economy. The Review of Economics and Statistics 33, 318-328.

ISARD, W. (1953). Regional commodity balances and interregional commodity flows. American Economic Review 43, 167-180.

JAYNES, E.T. (1957). Information theory and statistical mechanics. Physical Review 106, 620-630.

JAYNES, E.T. (1979). Where do we stand on maximum entropy? dans The Maximum Entropy Formalism, R.D. Levine and Myron Tribus, eds., M.I.T. Press.

LEONTIEFF, W. (1953). Interregional trade, in Studies in the Structure of the American Economy, W. Leontieff, ed., Oxford University Press, New York, 93-115.

LEONTIEFF, W. and STROUT, A. (1963): Multiregional input-output analysis, dans Structural Interdependence and Economic Development, Tibor Barna, ed. St. Martin's Press, Londres, 119-150.

LOS, M. (1980). Conception d'un modèle de prévision des flux interrégionaux de marchandises pour le Canada. Cahier 80-07, Département de sciences économiques, Université de Montréal.

MACGILL, S.M. (1977). Activity-Commodity Spatial Interaction Models and Related Applications. Ph.D, Thesis, School of Geography, University of Leeds, Leeds, U.K.

MACGILL, S.M. (1978). Rectangular input-output tables, multiplier analysis and entropy maximizing principles: A new methodology. Regional Science and Urban Economics 8, No. 4, 355-370.

Ministère de l'Expansion économique régionale, Division d'Analyse du Développement économique (1976). An Interprovincial Input-Output Model - Version III.

Ministère de l'Expansion économique régionale, Division d'Analyse du Développement économique (1977). Employment and Occupational Impacts Using the Version III Interprovincial Input-Output Model.

MOSES, L.N. (1955). The stability of interregional trading patterns and input-output analysis. American Economic Review 65, 803-832.

RIEFLER, R.F. (1973). Interregional input-output: A state of the arts survey, dans Judge, G.G. et Takayama, T., eds. Studies in Economic Planning over Space and Time, North-Holland, 133-162.

SNICKARS, F. (1979). Construction of interregional input-output tables by efficient information adding, dans Exploratory and Explanatory Statistical Analysis of Spatial Data, Bartels et Kettellaper, eds., Nijhoff, 73-111.

SNICKARS, F. et WEIBULL, J.W. (1977). A minimum information principle. Regional Science and Urban Economics 7, 137-168.

Statistique Canada, Division de l'Analyse structurelle (1976). User's Guide to Statistics Canada Structural Economic Models.

THEIL, H. (1967). Economics and Information Theory. North-Holland.

VAN PEETERSSEN, A. (1979). Procédure d'élaboration d'un tableau d'input-output interrégional pour le Canada. Rapport de recherche No 79-07, Ecole des Hautes Etudes Commerciales, Montréal.

WILSON, A.G. (1970). Entropy in Urban and Regional Modelling. Pion, Londres, Royaume-Uni.